

Criação e avaliação preliminar de um questionário padronizado para aferir experiência de usuário em realidade virtual imersiva

Jean Felipe Patikowski Cheiran^{a,b}, Denise Ruschel Bandeira^c, Marcelo Soares Pimenta^b

^a Campus Alegrete – Universidade Federal do Pampa (UNIPAMPA)

^b Instituto de Informática e ^c Instituto de Psicologia – Universidade Federal do Rio Grande do Sul (UFRGS)

jeancheiran@unipampa.edu.br, bandeira@ufrgs.br, mpimenta@inf.ufrgs.br

Resumo: A popularização da Realidade Virtual (RV) ampliou a necessidade de avaliação da Experiência do Usuário (UX) de forma rápida, barata e confiável. Assim, desenvolvemos um questionário padronizado em português brasileiro para avaliação de UX em RV imersiva, estabelecendo um modelo teórico de UX em RV, criando itens a partir da análise de conteúdo de outros instrumentos e avaliando os itens em um estudo com especialistas. A versão preliminar do questionário possui 47 itens divididos entre os componentes de usabilidade, estética visual clássica, estímulo, presença, *VR sickness* e emoções. Esses itens serão ajustados após um experimento piloto e as propriedades psicométricas do questionário final serão avaliadas em um experimento de longo prazo.

Palavras-chave: experiência de usuário; realidade virtual; questionário padronizado.

1. Introdução

A Realidade Virtual (RV), definida como um ambiente artificial interativo que é experienciado através de estímulos sensoriais fornecidos por um computador [20], tem passado por um período de popularização nos últimos anos. A comercialização de dispositivos imersivos de RV de preço acessível e os amplos investimentos em metaverso estão entre os fatores que impulsionam a criação de novos ambientes de RV para pesquisa, lazer, trabalho etc.

Tendo em vista que uma parte fundamental do desenvolvimento de sistemas interativos é a avaliação de Interação Humano-Computador (IHC) [1], a RV exige a consideração de aspectos particulares que extrapolam a qualidade em uso como, por exemplo, a sensação de presença (sensação de estar em um lugar diferente daquele em que o corpo físico está, sendo uma reação psicológica e subjetiva a ambientes imersivos [6]) e *VR sickness* (conjunto de sintomas desagradáveis similares a *motion sickness* e decorrentes do uso de Realidade Virtual, incluindo desorientação e enjoo [3]). Essa perspectiva mais ampla nos leva à noção de Experiência de Usuário (UX, do inglês *User eXperience*), que engloba todas as percepções e respostas do usuário que resultam do uso ou da expectativa de uso de um produto [8].

Embora o debate sobre abordagens quantitativas e qualitativas seja intenso e sua escolha dependa dos objetivos e do contexto da avaliação [1], questionários são amplamente usados como método de estimação de UX [13,19]. A possibilidade de avaliação da qualidade psicométrica dos questionários do ponto de vista de validade, fidedignidade e sensibilidade é um fator determinante na criação de questionários padronizados, ou seja, instrumentos projetados para uso recorrente, com itens em ordem e formato específicos, incluindo instruções de aplicação e interpretação de resultados [14].

Mesmo que questionários sejam uma ferramenta adequada para muitos cenários de avaliação de UX, é possível identificar três problemas que afetam

profissionais e pesquisadores brasileiros que queiram utilizá-los na avaliação de aplicações em RV: (a) a menor precisão nos resultados obtidos por meio de questionários *ad hoc*, (b) a falta de questionários padronizados traduzidos para português brasileiro e adequadamente validados e (c) as fragilidades dos questionários padronizados existentes para avaliar UX no contexto de RV de uma forma holística.

Dessa forma, o objetivo desse trabalho é desenvolver um questionário padronizado em português brasileiro para avaliação de Experiência de Usuário em Realidade Virtual imersiva com base na Teoria Clássica dos Testes. Para isso, propusemos um modelo teórico de UX em RV, identificamos questionários significativos para a avaliação de UX em RV (incluindo seus componentes), desenvolvemos uma coleção de itens a partir de uma análise de conteúdo dos principais instrumentos identificados, e montamos uma versão preliminar do questionário com base em uma avaliação de qualidade e pertinência dos itens com especialistas.

2. Trabalhos Relacionados

Diversos questionários padronizados utilizados nos processos de avaliação já foram mapeados na literatura científica de IHC [13,19] e RV [3,6,9]. A maior parte dos instrumentos identificados aborda uma UX genérica e agnóstica em relação ao domínio e à tecnologia ou aborda componentes particulares de UX como, por exemplo, usabilidade, engajamento, presença etc.

Considerando que a UX é composta por inúmeros aspectos determinados pelo contexto [16], o uso de questionários padronizados de UX genérica para avaliar experiências em RV frequentemente é insuficiente, porque se assume que a influência de componentes específicos de RV nos resultados é irrelevante ou que os componentes específicos de RV (não medidos) têm correlação significativa com outros componentes (medidos).

Por outro lado, a avaliação de UX em RV utilizando múltiplos questionários pode levar a outros problemas causados, por exemplo, por formatos distintos dos itens

e por incertezas sobre como um escore total de UX poderia ser calculado a partir de escores parciais [17].

Poucos questionários padronizados para avaliação de UX em RV foram identificados nessa pesquisa, sendo os mais abrangentes o Virtual Experience Test (VET) [5], o User Experience in Immersive Virtual Environments questionnaire (UXIVE) [17] e um questionário sem nome de Cheng et al. [4]. Contudo, algumas fragilidades nos processos de construção e validação desses instrumentos (e.g., falta de evidências de validade ou estimativas de fidedignidade, uso de técnicas pouco adequadas na análise de dimensionalidade, e ausência de instruções para calcular os escores) comprometem a interpretação de seus resultados.

Finalmente, cabe destacar que a quantidade de instrumentos de avaliação de UX ou de seus componentes em português é bastante pequena [11], tendo em vista que processos sistemáticos de tradução e adaptação precisam ser acompanhados de novas evidências de validade e estimativas de fidedignidade que demonstrem que as propriedades do instrumento não foram modificadas [7].

3. Metodologia e Resultados Preliminares

Essa pesquisa se caracteriza como um estudo correlacional, quantitativo e experimental. A Figura 1 apresenta as atividades já executadas nessa pesquisa, cobrindo etapas de fundamentação teórica do trabalho, planejamento e criação do questionário preliminar.

Na **identificação de modelos teóricos**, encontramos apenas dois modelos que incluem componentes de RV em suas estruturas: um modelo que investiga a UX em uma perspectiva de *flow* [4] e um modelo que descreve UX de um ponto de vista holístico [17]. Todavia, esses modelos não se mostraram adequados para uso na construção do questionário, pois (a) a estrutura proposta

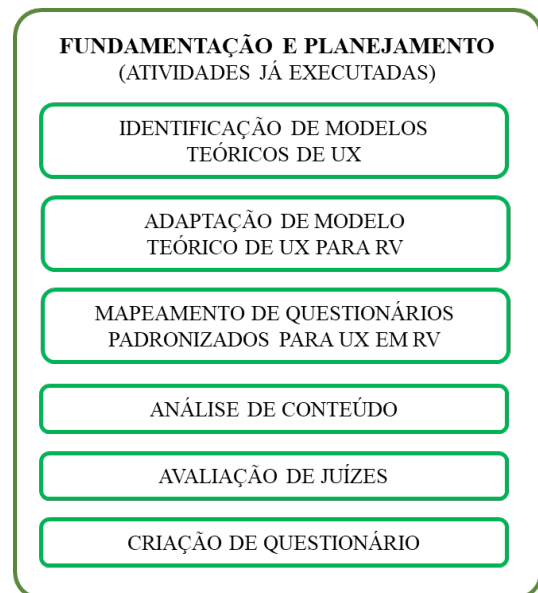


Figura 1. Atividades executadas em ordem cronológica.

por [4] não inclui elementos importantes de UX por se concentrar na experiência de *flow*, (b) o modelo de [17] é incompatível com resultados de outros estudos sobre a relação entre componentes da UX e (c) os dois modelos são parcialmente contraditórios entre si (e.g., a direção de causalidade entre *flow* e presença é oposta nos dois estudos). Embora apresentem algumas fragilidades, esses modelos contribuem significativamente para a identificação de componentes de UX e suas relações no contexto de RV, permitindo ampliar a investigação da estrutura mais adequada.

A partir dos modelos encontrados [4,17] e de uma revisão de modelos de UX independentes de domínio e tecnologia, realizamos a **adaptação do modelo** de

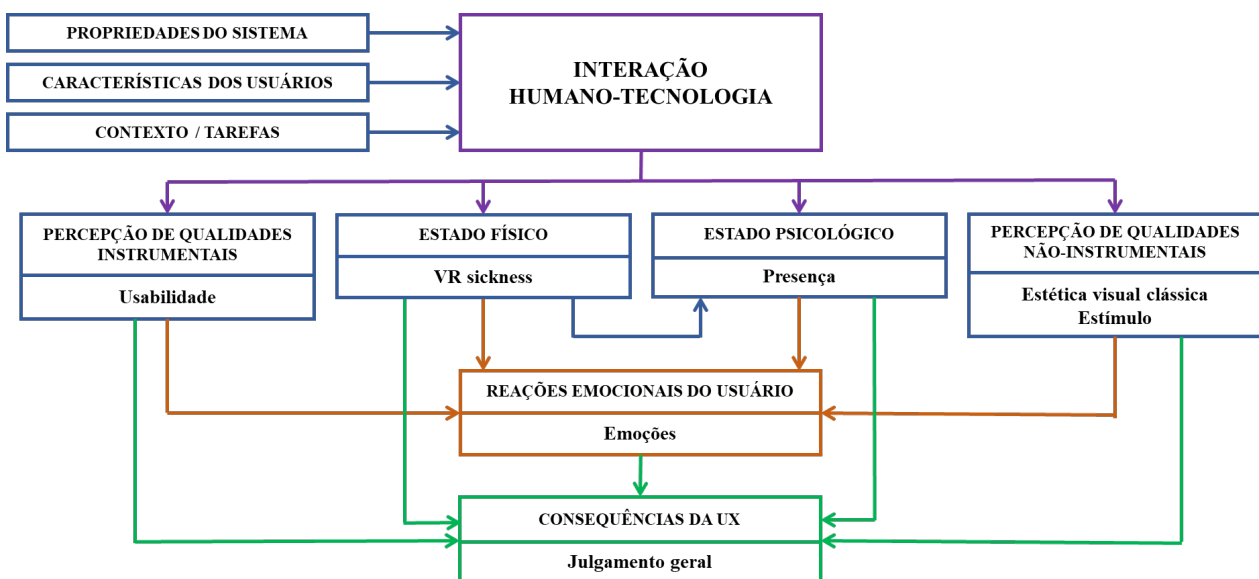


Figura 2. Modelo teórico de UX para RV (não confirmado) adaptado de [18].

Componentes de Experiência de Usuário (CUE) [18] para uso nessa pesquisa. A escolha do CUE se deu por sua coerência estrutural, sua flexibilidade para alterações e sua ampla validação experimental [18]. A Figura 2 ilustra o modelo (ainda não confirmado) que introduz componentes de *VR sickness* e presença, além de reestruturar alguns componentes do CUE original.

A identificação dos principais componentes de UX avaliados em RV extrapola os modelos teóricos, pois a organização dos itens e das subdimensões (modelos de medição) dos questionários existentes também contribuiu na compreensão do construto avaliado. Assim, um **mapeamento sistemático** foi conduzido para identificar questionários para aferir UX em ambientes de RV cujos estudos relatassem evidências de validade e estimativas de fidedignidade [7]. A partir de 360 artigos obtidos em seis bases científicas, 20 artigos foram selecionados. Achados importantes desse mapeamento incluem a falta da descrição de procedimentos para formar um escore na maior parte dos artigos e o uso de RV imersiva e jogos nos experimentos da maior parte dos estudos.

Os 20 questionários obtidos no mapeamento sistemático compuseram a coleção inicial de documentos para **análise de conteúdo** [12]. Também foram incluídos outros 29 questionários identificados em revisões de outros autores [3,6,9,13,19], posto que estavam entre os mais utilizados em pesquisas na área de RV. A estimação de uso dos questionários foi feita a partir de citações em bases acadêmicas. No total, 49 questionários foram examinados em um processo de agrupamento de itens similares nas dimensões inicialmente exploradas: usabilidade (dividida em facilidade de aprendizado, eficiência e eficácia), estética visual clássica, estética expressiva (dividida em estímulo e novidade), presença (dividida em imersão, coerência, ilusão de lugar e ilusão de plausibilidade), *VR sickness* e emoções. Cada uma das 149 categorias (agrupamentos) que emergiram a partir de itens similares foi então compilada em um único item que representasse a ideia geral da categoria.

A coleção de 149 itens foi **avaliada por juízes** especialistas (doutores) das áreas de Experiência de Usuário (2 juízes) e Realidade Virtual (2 juízes). Essa avaliação tem objetivo de identificar problemas de qualidade e a pertinência dos itens para medir os componentes investigados [7]. Ela envolveu classificar a qualidade dos itens em relação à forma de escrita (categorias: ruim, regular ou bom), especificar a pertinência dos itens para medir UX e seus componentes (categorias: muito dispensável, dispensável, incerto, importante ou muito importante), adicionar comentários sobre os itens e incluir itens novos. O processo de avaliação foi remoto, síncrono e individual. Destaca-se que 21 itens novos (incluindo variações dos itens existentes) foram propostos pelos juízes durante as avaliações, totalizando 170 itens candidatos.

A **criação do questionário** baseou-se nas avaliações dos juízes e em seus comentários. Inicialmente, 3 componentes de UX que faziam parte do modelo originalmente proposto foram removidos (imersão, coerência e novidade), assim como todos os itens

relacionados a eles. Os demais itens foram analisados individualmente a partir das classificações e dos comentários associados. Ao todo, 106 itens dos 149 resultantes da análise de conteúdo e 17 itens dos 21 propostos pelos juízes foram descartados. O questionário em sua primeira versão (Disponível em <https://doi.org/10.6084/m9.figshare.20520741>) possui 47 itens divididos nos seguintes componentes:

- Usabilidade – facilidade de aprendizado (5 itens)
- Usabilidade – eficiência (5 itens)
- Usabilidade – eficácia (6 itens)
- Presença – ilusão de lugar (5 itens)
- Presença – ilusão de plausibilidade (5 itens)
- Estética visual clássica (5 itens)
- Estímulo (5 itens)
- Emoções (5 itens)
- VR sickness (6 itens)

O questionário é composto por itens no estilo Likert: afirmações com uma faixa de resposta simétrica, etiquetada e numerada na qual o usuário indica sua concordância. A faixa de resposta contém 7 pontos para viabilizar etiquetas coerentes e não ambíguas. A Figura 3 mostra as atividades pendentes de experimentação, avaliação e análise dos resultados.

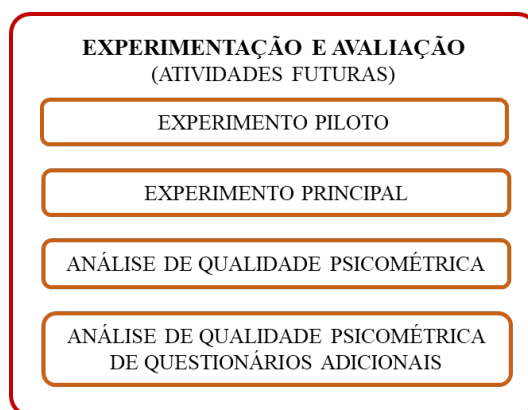


Figura 3. Atividades pendentes em ordem cronológica.

Um **experimento piloto** está sendo usado no lugar de grupos focais [7], uma vez que compreender muitos itens do questionário depende do uso de RV. Os participantes, depois de utilizarem aplicativos através de óculos de RV (Oculus Quest 2) e responderem ao questionário após cada uso, serão inquiridos sobre a clareza dos itens do questionário. Os aplicativos incluem um ambiente tutorial, um jogo de ritmo e um aplicativo de pintura.

Após a realização de ajustes no questionário, o **experimento principal** será conduzido de forma praticamente idêntica ao experimento piloto, embora nenhum *feedback* sobre a qualidade dos itens seja coletado. Além disso, questionários adicionais [2,10,15] serão respondidos por participante. A quantidade estimada de participantes é 150. A técnica *latin squares*

é usada para que participantes usem aplicativos e respondam a questionários em diferentes ordens.

A **análise de qualidade psicométrica** se concentrará em estimativas de fidedignidade (coeficientes ômega) e evidências de validade [7] baseadas no conteúdo do teste (argumentação sobre itens baseada em modelos teóricos e questionários existentes, avaliação por juízes, avaliação em estudo piloto), na estrutura interna (análises fatoriais, correlação ítem-total, análise de invariância da medida) e em relações com outras variáveis (correlação entre dimensões do questionário e outros questionários existentes). A investigação sobre o modelo estrutural será realizada através de análise de caminhos (*path analysis*). Finalmente, a **análise das propriedades psicométricas de outros questionários** usados [2,10,15] ocorrerá de forma similar àquela do questionário proposto.

4. Considerações Preliminares

Atualmente, essa pesquisa se encontra na realização do estudo piloto cujos resultados ainda estão sob análise e, portanto, não serão discutidos nesse artigo.

A principal contribuição desse trabalho para desenvolvedores de aplicações em RV no Brasil é o questionário padronizado que permite uma estimação da experiência subjetiva dos usuários dessas aplicações. Adicionalmente, espera-se oferecer para pesquisadores brasileiros um modelo teórico de UX em RV que viabilize novas investigações (incluindo a extensão do questionário desenvolvido) e, também, descrever um processo detalhado de criação e avaliação psicométrica de um questionário padronizado.

Referências

- [1] Barbosa, S. D. J.; Silva, B. S.; Silveira, M. S.; Gasparini, I.; Darin, T.; Barbosa, G. D. J. (2021). *Interação Humano-Computador e Experiência do Usuário*. 1ª ed. Autopublicação.
- [2] Carvalho, M. R.; Costa, R. T.; Nardi, A. E. (2011). Simulator Sickness Questionnaire: tradução e adaptação transcultural. *Jornal Brasileiro de Psiquiatria* 60(4). 0–5. DOI: [10.1590/S0047-20852011000400003](https://doi.org/10.1590/S0047-20852011000400003)
- [3] Chang, E.; Kim, H. T.; Yoo, B. (2020). Virtual Reality Sickness: A Review of Causes and Measurements. *Int. Journal of Human-Computer Interaction*, 36(17). 1658–1682. DOI: [10.1080/10447318.2020.1778351](https://doi.org/10.1080/10447318.2020.1778351)
- [4] Cheng, L. K.; Chieng, M. H.; Chieng, W. H. (2014). Measuring virtual experience in a three-dimensional virtual reality interactive simulator environment: A structural equation modeling approach. *Virtual Reality* 18(3). 173–188. DOI: [10.1007/s10055-014-0244-2](https://doi.org/10.1007/s10055-014-0244-2)
- [5] Chertoff, D. B.; Goldiez, B.; LaViola, J. J. (2010). Virtual experience test: A virtual environment evaluation questionnaire. In *IEEE Virtual Reality Conference*. IEEE. DOI: [10.1109/VR.2010.5444804](https://doi.org/10.1109/VR.2010.5444804)
- [6] Grassini, S.; Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology* 11: 349. DOI: [10.3389/fpsyg.2020.00349](https://doi.org/10.3389/fpsyg.2020.00349)
- [7] Hutz, C. S.; Bandeira, D. R.; Trentini, C. M. (editores) (2015). *Psicometria*. Artmed.
- [8] ISO 9241-210:2019 - ergonomics of human-system interaction - part 210: Human-centred design for interactive systems. <https://www.iso.org/standard/77520.html> Acesso 19/08/22
- [9] Kim, Y. M.; Rhiu, I.; Yun, M. H. (2020). A Systematic Review of a Virtual Reality System from the Perspective of User Experience. *International Journal of Human-Computer Interaction* 36(10). 893–910. DOI: [10.1080/10447318.2019.1699746](https://doi.org/10.1080/10447318.2019.1699746)
- [10] Martins, A. I.; Rosa, A. F.; Queirós, A.; Silva, A.; Rocha, N. P. (2015). European Portuguese validation of the System Usability Scale (SUS). *Procedia - Procedia Computer Science*, 67. 293–300. DOI: [10.1016/j.procs.2015.09.273](https://doi.org/10.1016/j.procs.2015.09.273)
- [11] Melo, B.; Darin, T. (2019). Scope and definition of user experience in Brazil: A survey to explore community's perspectives. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*. ACM. DOI: [10.1145/3357155.3358444](https://doi.org/10.1145/3357155.3358444)
- [12] Moraes, R. (1999). Análise de conteúdo. *Revista Educação*, 22(37). 7–32.
- [13] Rivero, L.; Conte, T. (2017). A systematic mapping study on research contributions on UX evaluation technologies. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*. ACM. DOI: [10.1145/3160504.3160512](https://doi.org/10.1145/3160504.3160512)
- [14] Sauro, J.; Lewis, J. (2016). *Quantifying the User Experience: Practical Statistics for User Research*. 2nd ed. Elsevier.
- [15] Silva, G. R.; Donat, J. C.; Rigoli, M. M.; Oliveira, F. R.; Kristensen, C. H. (2016). A questionnaire for measuring presence in virtual environments: factor analysis of the presence questionnaire and adaptation into Brazilian Portuguese. *Virtual Reality* 20(4). 237–242. DOI: [10.1007/s10055-016-0295-7](https://doi.org/10.1007/s10055-016-0295-7)
- [16] Schrepp, M.; Thomaschewski, J. (2019). Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia and Artificial Intelligence* 5(7). 88–95. DOI: [10.9781/ijimai.2019.06.006](https://doi.org/10.9781/ijimai.2019.06.006)
- [17] Tcha-Tokey, K.; Loup-Escande, E.; Christmann, O.; Canac, G.; Farin, F.; Richir, S. (2018). Towards a user experience model in immersive virtual environments. *Advances in Human-Computer Interaction* 2018. 1–10. DOI: [10.1155/2018/7827286](https://doi.org/10.1155/2018/7827286)
- [18] Thüring, M.; Mahlke, S. (2007). Usability, aesthetics and emotions in human-technology interaction. *International Journal of Psychology* 42(4). 253–264. DOI: [10.1080/00207590701396674](https://doi.org/10.1080/00207590701396674)
- [19] Vermeeren, A.; Law, E.; Roto, V.; Obrist, M.; Hoonhout, J.; Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: Current state and development needs. In *Proceedings of the NordiCHI*. ACM. DOI: [10.1145/1868914.1868973](https://doi.org/10.1145/1868914.1868973)
- [20] Virtual Reality. <https://www.merriam-webster.com/dictionary/virtualreality>. Acesso: 19/08/2022